Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

MNRAS **437,** 3133–3143 (2014)
Advance Access publication 2013 December 9

doi:10.1093/mnras/stt2078

Trawling for transits in a sea of noise: a search for exoplanets by analysis of WASP optical light curves and follow-up (SEAWOLF)

E. Gaidos,[1]★ D. R. Anderson,[2] S. Lépine,[3] K. D. Colón,[1] G. Maravelias,[4] N. Narita,[5] E. Chang,[1] J. Beyer,[1] A. Fukui,[6] J. D. Armstrong,[7] A. Zezas,[4] B. J. Fulton,[7,8] A. W. Mann,[7] R. G. West[9] and F. Faedi[9]

[1]*Department of Geology & Geophysics, University of Hawaii at Mānoa, Honolulu, HI 96822, USA*
[2]*Astrophysics Group, Keele University, Staffordshire ST5 5BG, UK*
[3]*Department of Astrophysics, American Museum of Natural History, New York, NY 10024, USA*
[4]*Physics Department & Institute of Theoretical & Computational Physics, University of Crete, GR-71003 Heraklion, Crete, Greece*
[5]*National Astronomical Observatory of Japan, 2-21-1 Osawa, Mitaka, Tokyo 181-8588, Japan*
[6]*Okayama Astrophysical Observatory, National Astronomical Observatory of Japan, Asakuchi, Okayama 719-0232, Japan*
[7]*Institute for Astronomy, University of Hawaii at Mānoa, Honolulu, HI 96822, USA*
[8]*Las Cumbres Observatory Global Telescope Network, Goleta, CA 93117, USA*
[9]*Department of Physics, University of Warwick, Coventry CV4 7AL, UK*

Accepted 2013 October 23. Received 2013 August 9; in original form 2013 July 10

ABSTRACT

Studies of transiting Neptune-sized planets orbiting close to nearby bright stars can inform theories of planet formation because mass and radius and therefore mean density can be accurately estimated and compared with interior models. The distribution of such planets with stellar mass and orbital period relative to their Jovian-mass counterparts can test scenarios of orbital migration, and whether 'hot' (period <10 d) Neptunes evolved from 'hot' Jupiters as a result of mass loss. We searched 1763 late K and early M dwarf stars for transiting Neptunes by analysing photometry from the Wide Angle Search for Planets and obtaining high-precision ($\leq 10^{-3}$) follow-up photometry of stars with candidate transit signals. One star in our sample (GJ 436) hosts a previously reported hot Neptune. We identified 92 candidate signals among 80 other stars and carried out 148 observations of predicted candidate transits with 1–2 m telescopes. Data on 70 WASP signals rules out transits for 39 of them; 28 other signals are ambiguous and/or require more data. Three systems have transit-like events in follow-up photometry and we plan additional follow-up observations. On the basis of no confirmed detections in our survey, we place an upper limit of 10.2 per cent on the occurrence of hot Neptunes around late K and early M dwarfs (95 per cent confidence). A single confirmed detection would translate to an occurrence of 5.3 ± 4.4 per cent. The latter figure is similar to that from Doppler surveys, suggesting that GJ 436b may be the only transiting hot Neptune in our sample. Our analysis of *Kepler* data for similar but more distant late-type dwarfs yields an occurrence of 0.32 ± 0.21 per cent. Depending on which occurrence is applicable, we estimate that the Next Generation Transit Survey will discover either ∼60 or ∼1000 hot Neptunes around late K- and early M-type dwarfs.

Key words: techniques: photometric – surveys – planets and satellites: detection – planets and satellites: formation – planets and satellites: general – stars: late-type.

1 INTRODUCTION

Not all exoplanets are detected equally. A planet that transits its host star has greater scientific value because its radius can be determined and, because the orbital inclination is known, the geometric ambiguity in Doppler estimation of the planet mass is removed. Spectroscopy of the star during a transit can reveal absorption or scattering by the planet's atmosphere, if it has one. The planet can also be occulted by the star, permitting differential measurement of the planet's reflected or emitted flux. These observations can determine the planet's albedo and/or constrain the efficiency with which heat is carried around the planet by rotation or atmospheric circulation.

★ E-mail: gaidos@hawaii.edu

The most productive tool for detecting transiting planets has been the *Kepler* space telescope (Borucki et al. 2010), data from which have yielded more than 2000 confirmed or candidate discoveries. However, most of the systems discovered by *Kepler*, as well as those of the COnvection ROtation et Transits planétaires (CoRoT) satellite (Carone et al. 2012), are faint ($V \sim 15$), making follow-up observations difficult. Many of these host stars are at kpc distances and well above the Galactic plane, and may belong to an older, more metal-poor population distinct from the solar neighbourhood and perhaps hosting a different distribution of planets.

Ground-based surveys such as the Wide Angle Search for Planets (WASP; Pollacco et al. 2006) and the Hungarian Automated Telescope Network (HatNET; Bakos et al. 2011) have discovered numerous giant planets transiting brighter, nearby stars.[1] Transiting geometries are uncommon and such surveys must monitor many stars over large portions of the sky. Because of the trade-off between field of view and telescope aperture, these surveys are limited to the brightest stars and, due to Malmquist bias, biased towards more luminous ones. The sensitivity of such surveys to smaller (non-gas-giant) planets is limited by correlated photometric error or 'red' noise which does not decrease with the square root of the number of observations (Pont, Zucker & Queloz 2006; Smith et al. 2007). Earth's rotation means that surveys performed from a single site have restricted observing windows and are only efficient at detecting planets on short-period orbits ($\leq$10 d). For these reasons, ground-based surveys have been most successful at detecting giant planets on close orbits around F and G stars.[2]

M dwarf stars have less than half the radius of their solar-type cousins, permitting the detection of concomitantly smaller planets for a given photometric sensitivity. Although such stars tend to be fainter and observed with poorer photometric precision, the net balance of these two effects can still favour cooler stars: this calculus motivates the MEarth transit survey for planets as small as Earth around late M-type dwarfs (Charbonneau et al. 2009; Berta et al. 2012).

K- and early M-type dwarfs represent an intermediate region of discovery space for transiting planet surveys. While ground-based detection of Earth-sized planets around such stars is not feasible, it is possible to detect Neptune- or even super-Earth-sized companions, at least on close-in orbits. Indeed, HAT-P-11b ($4.3R_\oplus$, $P = 4.89$ d) transits a K4 dwarf, and HAT-P-26b ($6.3R_\oplus$, $P = 4.23$ d) orbits a K1 dwarf (Hartman et al. 2011).[3]

The occurrence and properties of short-period or 'hot' Neptunes are of considerable theoretical interest. Attempts to explain an apparent correlation between the occurrence of giant planets and stellar mass also predict an inverse relation with elevated numbers of Neptunes (i.e. 'failed Jupiters') around low-mass stars (Laughlin, Bodenheimer & Adams 2004). Hot Neptunes could form by accretion of rocky/icy planetesimals beyond the snowline and subsequent migration to the inner edge of the protoplanetary disc (Mordasini, Alibert & Benz 2009). However, McNeil & Nelson (2010) find that this scenario cannot explain the observed size distribution of close-in planets. Alternatively, planetesimals or protoplanets could migrate first, followed by accretion in place (Brunini & Cionco 2005; Hansen & Murray 2012). Finally, evaporation of mass from close-in giant planets has been proposed as an alternative formation mechanism for hot Neptunes (Baraffe et al. 2005; Boué et al. 2012). These three different pathways predict objects that are enriched in ice, rock and gas, respectively. Hot Neptunes may be especially useful to test models of planet formation because both mass and radius can be accurately measured (by Doppler and transit, respectively) and these parameters are informative about the relative amounts of rock, ice and gas in the planet (Rogers et al. 2011). In contrast, the masses of Earth-sized planets are too small to accurately measure and the radii of Jupiter-sized planets are insensitive to mass due to support by electron degeneracy pressure.

We used data from the WASP survey to search for short-period Neptunes around a sample of low-mass stars (the SEAWOLF survey). Because this search pushes the envelope of WASP performance, we adopted a multistage search strategy.

(i) We selected late K- and early M-type dwarf stars observed by the WASP survey; in principle, smaller planets should be detectable around these smaller stars.

(ii) We identified candidate transit signals, relaxing the signal-to-noise ratio criterion for initial selection. This potentially includes not only smaller transit signals, but also large numbers of false positives.

(iii) We predicted candidate transits using the WASP-generated ephemerides and screened these with precision photometry obtained at 1–2 m telescopes.

We describe the WASP data and our follow-up observations and reduction in Section 2, and our catalogue of candidate transiting systems and the results of the follow-up program in Section 3. We place limits on the occurrence of hot Neptunes around stars in our sample in Section 4, and discuss the implications for theory as well as prospects for future transiting planet surveys in Section 5.

2 OBSERVATIONS AND METHODS

2.1 Sample construction and stellar parameters

For our search sample we identified late-type (K4 to M4) dwarf stars in the inaugural (2004) fields of the WASP-North survey (Christian et al. 2006). We chose stars from the SUPERBLINK proper motion catalogue (Lépine & Shara 2005) with optical-to-infrared colours $V - J > 2$ consistent with late K- and M-type stars, and reduced proper motions $H_J \equiv J + 5\log\mu + 5$ (a proxy for absolute magnitude) that place them on the dwarf colour–magnitude locus, thus excluding K and M giants (Lépine & Gaidos 2011). We restricted the sample to $V < 14$ because at fainter magnitudes the number of background stars with $\Delta m < 5$ falling within the same WASP photometric aperture significantly exceeds one. Such stars could produce false positives if they are eclipsing binaries. We also imposed a $J < 10$ cut to retain those stars for which high-precision, high cadence (few minute) photometry in the near-infrared (z or *JHK* passbands) could be performed on 1–2 m telescopes. Based on parallaxes (astrometric wherever available, photometric otherwise), the most distant stars in our survey are at $\approx$100 pc. The closest star is Laland 21185, only 2.5 pc away.

To estimate the properties of these stars, we adopted the empirical relations between $V - J$ colour, effective temperature $T_{\rm eff}$, stellar radius R_*, and stellar mass M_* for solar-metallicity K and M stars

[1] Transiting planets have also been identified by screening planetary systems detected by the Doppler method. The first example (HD 209458) was found this way (Charbonneau & Brown 2000; Henry et al. 2000), but this approach is limited by the pace of Doppler surveys and the small geometric probability that a planet will transit.

[2] Wide-field surveys must also contend with a high false positive rate by blends of bright stars with fainter eclipsing binaries.

[3] Two other Neptune-sized planets, both around early-type M dwarfs, were detected first by Doppler, then later found to transit: GJ 436b (Gillon et al. 2007) and GJ 3740b (Bonfils et al. 2012).

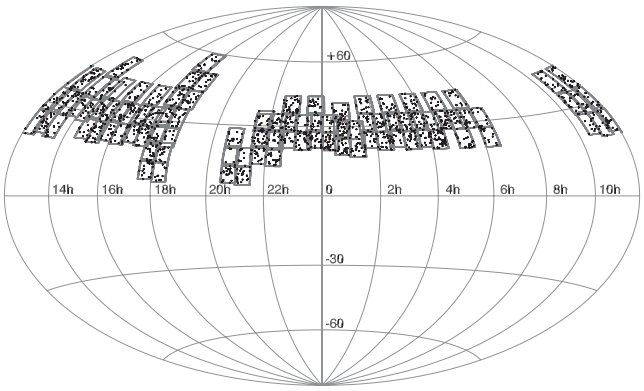

Figure 1. Locations of the 2004 inaugural fields of the WASP-North survey and our selected SUPERBLINK K and M dwarf stars.

in Boyajian et al. (2012).[4] According to these relations, $V - J = 2$ corresponds to $R_* \approx 0.71\ \mathrm{R}_\odot$, $T_{\rm eff} \approx 4550$ K, and a spectral subtype of K4 (Cox 2000). The coolest stars in our sample have $V - J \geq 4.5$ and should have M4 spectral types, with $R_* \approx 0.25\ \mathrm{R}_\odot$ and $T_{\rm eff} \approx 3300$ K. The reddest star ($V - J = 5.39$) is the M4.5 dwarf GJ 3839.

2.2 WASP observations and sources

We identified 1849 SUPERBLINK stars satisfying our criteria in the inaugural (2004) fields of the WASP-North survey. These 102 fields cover 4750 deg^2 at declinations between $+4\overset{\circ}{.}9$ and $+59\overset{\circ}{.}3$. (Fig. 1). Stars were matched with sources generated by photometering WASP images with a circular aperture of radius 3.5 pixels (48 arcsec). 1763 WASP sources were matched to our selected SUPERBLINK stars; the median angular separation is 0.22 arcsec and the 95th percentile separation is 3.4 arcsec. 46 WASP sources were each matched to two SUPERBLINK stars. Of the 1763 matched sources, 1743 have more than 500 data points (the minimum required for light-curve analysis) and the median number of observations is 8160 (Fig. 2).

2.3 Light-curve analysis

WASP light curves were processed to correct for systematic errors (Tamuz, Mazeh & Zucker 2005) and remove trends (Kovács, Bakos & Noyes 2005). The latter step eliminates many artefacts with periods equal to rational multiples of 1 d. The light curves were then analysed with the HUNTER hybrid search algorithm which incorporates the box-least-squared algorithm (Kovács, Zucker & Mazeh 2002) and is described in Collier Cameron et al. (2006). HUNTER searched for transit-like signals with periods of 0.3–30 d. Four criteria were applied to these signals: (i) mean flux >3 microVegas ($m < 13.8$); (ii) periods not within 5 per cent of 1 or 0.5 d (see Section 2.4); (iii) signal detection efficiency >6 (Kovács et al. 2002); (iv) at least three candidate transits.

Up to five periodic signals were identified for each source satisfying these criteria, and a total of 4364 signals were identified among 1130 stars. HUNTER calculated the signal-to-red noise ratio (SRN) and $\Delta\chi^2$ parameter for each signal, where the latter is the decrease

[4] We used Two Micron All-Sky Survey (2MASS) J magnitudes while Boyajian et al. (2012) used Johnson J magnitudes. However, the CIT photometric system is closely related to the Johnson system and $J_{\rm CIT} \approx J_{\rm 2MASS} - 0.065(J-K)_{\rm 2MASS} + 0.038$ (Skrutskie et al. 2006). Since late K-early M dwarfs have $J - K \approx 0.8$, the difference in $V - J$ colour is only 0.014 mag and was ignored.

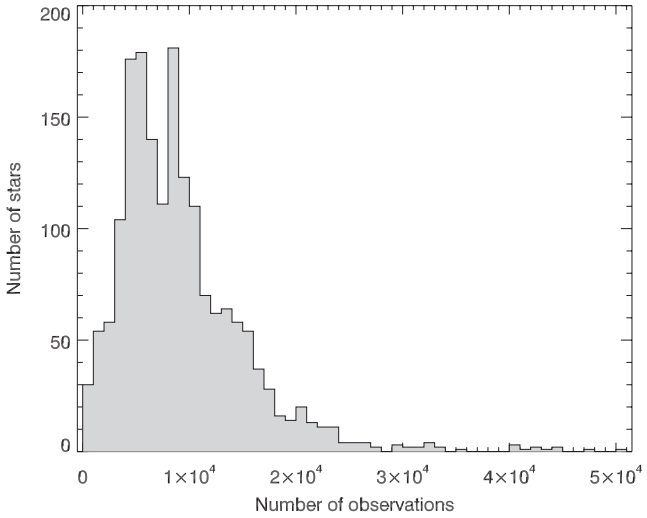

Figure 2. Distributions of number of WASP observations per star in our sample. The median number of observations per star is 8160.

in χ^2 provided by the best-fitting transit model relative to a constant light curve. In the case of pure white noise $\Delta\chi^2$ is the square of the signal-to-white-noise ratio (Collier Cameron et al. 2006). Thus SRN and $\Delta\chi^2$ allow us to select based on the significance of a signal with respect to both the red noise and white-noise properties of the data.

2.4 Selection of candidate transiting systems

We next applied cuts with period, SRN, $\Delta\chi^2$ and ellipsoidal signal-to-noise ratio (a measure of the continuous variation of the signal over the period) to the 4364 HUNTER-identified signals to screen artefacts and astrophysical false positives (i.e. close binaries). Because of Earth's rotation, observations from a single longitude like those of WASP-North can contain artefacts with periods near 1 d and integer ratios thereof. Furthermore, aliasing with the lunar cycle (29.5 d) produces a dispersion of a few per cent around each rational period. Based on the distribution of signals (mostly artefacts) generated when no detrending is performed (see above), we removed signals with periods below 1.1 d and within 5 per cent of 3/2, 2, 3 and 5 d (Fig. 3). (A peak at 4 d is not statistically significant.) There is also a peak in the period distribution of signals at 8/3 d. This peak appears significant and is apparently one of a series of undertones (multiples) of the strong artefact at 1/3 d, the harmonic closest to the duration of a summer night at the WASP-North site (and which is removed along with all other signals below 1.1 d). However, we did not a priori remove the 1/3 d peak.

The distribution of signals with SRN and $\Delta\chi^2$ is strongly concentrated at SRN ~4.5 and $\Delta\chi^2$ ~30 (Fig. 4). We assumed that these are nearly all artefacts or astrophysical false positives and that the clustering is a result of the selection criteria applied in Section 2.3. We retained signals with $(SRN > 6) \cup (SRN > 3 \cap \Delta\chi^2 > 50)$ (outside the hatched zone of Fig. 4). We also excluded signals with an ellipsoidal SNR >8: these are possible close binaries (Collier Cameron et al. 2006). The remaining 901 signals were further screened with the following criteria: (i) observations had to completely span, ingress to egress, at least three putative transits; (ii) the putative transit duration τ had to be within a factor of 2 of the value for a planet on a circular orbit with zero impact parameter around a star with radius 0.6 $\mathrm{R}_\odot$ (typical of a late K dwarf), i.e.

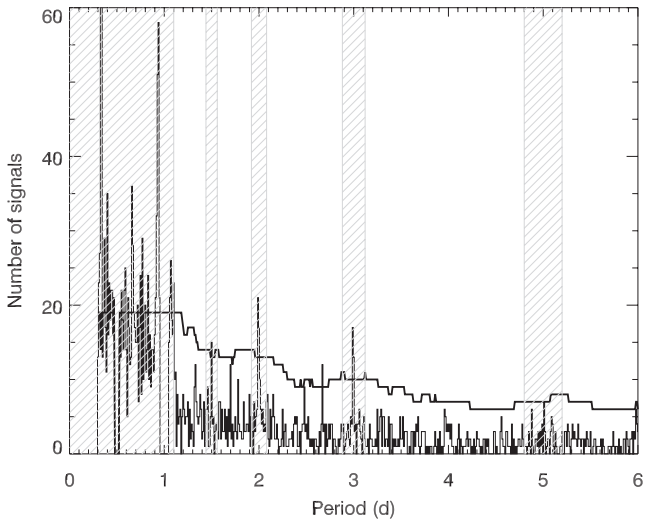

Figure 3. Period distribution of signals from the WASP HUNTER pipeline. The upper solid curve is a significance threshold ($p = 10^{-4}$) based on the Poisson statistics of a running mean ($n = 50$). Clusters of artefacts are present at rational multiples of 1 d. The hatched regions indicate exclusion zones around these periods and at <1.1 d; signals within these zones were rejected.

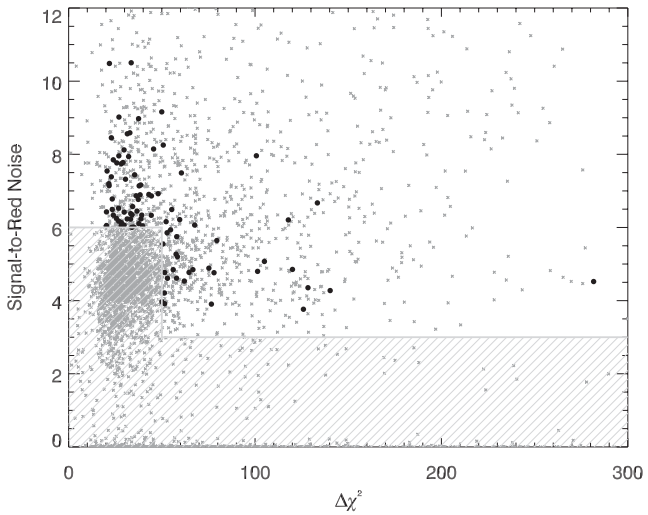

Figure 4. Signals from the WASP HUNTER pipeline; SRN versus $\Delta\chi^2$. Only signals outside the hatched zone were considered. The final candidate transiting systems are indicated by the large black points.

$\tau = 1\ \mathrm{hr}(P/1\mathrm{d})^{1/3}$, where P is the Keplerian period; and (iii) the putative signal could not obviously be an artefact produced by periodic gaps or changes in noise level in the data. This left 92 candidate signals from 80 stars.

2.5 Follow-up photometry

We used the ephemerides generated by the HUNTER pipeline to predict transits for our 92 candidates. The precision of the predicted transit centre depended mostly on the precision of the period determination, and was generally ±1 h. Follow-up photometry of some of these candidate transit events was obtained with 1–2 m ground-based telescopes. Details of the telescopes are reported in Table 1 and of the observations in Table 3. In general, we selected events to observe if the predicted transit centre occurred when the star was at an airmass below 1.7, and at least 3 h from sunset or sunrise. Ideally, we observed the entire transit window ($\pm 1\sigma$) as well as an hour before and after ingress/egress. However in many cases this was not possible. The minimum detectable transit depth δ_{d} and completeness C of these observations are calculated in Section 4.

Although the details of the observing strategy and data reduction varied with telescope and instrument, there were several commonalities:

Defocused imaging photometry. A telescope was grossly defocused to produce a 'doughnut'-shaped point spread function (PSF) several tens of pixels in diameter. Such 'doughnuts' are out-of-focus images of the primary mirror. Defocusing permitted a signal $\gg 1 \times 10^6$ e^{-1} to be acquired in each integration, reducing Poisson error to $<10^{-3}$. It also minimized error from image motion or changes in the distribution of the signal convolved with detector flat-fielding errors (e.g. Southworth et al. 2009; Mann, Gaidos & Aldering 2011). Circular aperture photometry was performed on the defocused images of the target and several comparison stars in multiple iterations. In each iteration the centroid of the stellar image within the aperture was computed and used as the aperture centre.

Optimized pointings. The signal from a star of interest must be divided by that from one or more comparison stars to remove variations in atmospheric transmission. The number and relative brightness of comparison stars limit the precision of ground-based photometry. We chose pointings that maximized the number of comparison stars similar in brightness to the target star. We also avoided rings in the flat-field due to dust particles near the focal plane. These can change between nights or even during observations, introducing flat-field error.

Comparison star selection. Each comparison star was compared with all the others to identify and exclude variables. A comparison signal was calculated from the weighted sum of the remaining

Table 1. Telescopes used to obtain follow-up observations.

Telescope/observatory	D (m)	Latitude	Longitude	Instrument(s)	Passband(s)	Observations[a]
McGraw-Hill/MDM	1.3	31.95173 N	111.61664 W	B4K/R4K/Nellie	Sloan r, DES-Z	66
Faulkes North/Haleakala	2.0	20.70701 N	156.25748 W	SpectraCam 4K	Pan-STARRS Z	29
Skinakas	1.3	35.21173 N	024.89893 E	Andor DZ436	Bessel R	24
OAO 188 cm	1.88	34.57716 N	133.59387 E	ISLE[b]	J	19
LCOGT/BOS	0.8	34.68750 N	120.03889 W	SBIG	Sloan i'	5
LCOGT/ELP	1.0	30.67143 N	104.02195 W	kb73	Sloan i	4
UH88/Mauna Kea	2.2	19.82303 N	155.46937 W	OPTIC	Sloan z	1

[a]Usable data of a candidate transit event ($C > 0$).
[b]Yanagisawa et al. (2006).

Table 2. Candidate transit systems identified in WASP data.

Name[a]	RA			Dec.			V	$V-J$	Period	Ephemeris	δ	Status[b]
	hh	mm	ss.s	dd	mm	ss			(d)	(BJD)	(10^{-3})	
00177+2100	0	17	43.2	21	00	05	12.4	2.45	5.319	3878.1780	4.6	?
00492+2003	0	49	17.2	20	03	45	10.8	2.28	8.063	3502.8216	3.3	?
01086+1714A	1	08	40.4	17	14	33	10.7	2.66	5.530	4177.7331	4.3	X
01086+1714B	1	08	40.4	17	14	33	10.7	2.66	3.743	4169.6813	3.2	?
01550+4035	1	55	01.0	40	35	06	13.7	3.91	8.343	3387.8129	20.3	?
01578+3130	1	57	50.0	31	30	41	12.2	2.25	4.177	4225.8173	4.1	X
01587+3515	1	58	43.6	35	15	28	13.7	4.01	7.819	4191.1928	9.7	X
02083+2919	2	08	18.3	29	19	59	12.5	2.93	7.214	3786.2246	8.0	X
02111+2707	2	11	11.2	27	07	34	12.7	3.36	10.560	3894.2242	8.8	N
02192+2456A	2	19	17.5	24	56	38	13.8	4.09	6.985	4022.5748	11.9	N

[a]ABC refer to multiple signals for the same star.
[b]X = ruled out, ? = ambiguous or insufficient data, A = candidate, N = not observed.
Table 2 is published in its entirety as a machine-readable table in the online version of this article and the Centre de Donées Strasbourg (CDS). A portion is shown here for guidance regarding its form and content.

reference stars, where the weights were chosen to minimize the rms of the normalized target light curve outside the predicted transit window.

Light curve detrending. We performed linear regressions of each normalized light curve with airmass, position of the centroid, and variance in the distribution of the target star signal over the PSF. The first was to remove second-order extinction effects due to differences in the spectra of target and reference stars (Mann et al. 2011). The second partly removes flat-field errors introduced when the defocused images move due to imperfect guiding or absence of guiding. The third compensates for any non-linearity in the response of the detector which would scale with the variance in the light distribution.

3 RESULTS

3.1 Candidate signals

The final catalogue of 92 candidate signals from 80 stars is presented in Table 2. The 'A', 'B' or 'C' designations indicate different signals from the same star. Based on the depths of the putative WASP transits and the estimated radii of these stars (see Section 2.1) the median transiting planet radius would be $\sim 4\ R_{\oplus}$, i.e. Neptune-sized. However, we expect that the large majority of these signals are artefacts or astrophysical false positives and not transiting planets. In Table 2 we report the status of each candidate based on the follow-up photometry acquired to date: N = not observed, ? = ambiguous or requires additional observations, X = eliminated, A = candidate transiting system. We have follow-up observations of 70 signals and we ruled out 39 signals and designated 28 as ambiguous or lacking sufficient data. In general, systems where the completeness C of our follow-up observations is >80 per cent (as calculated in Section 4.2), and no transit-like event was observed, were ruled out, and systems with one or more observations but where completeness was <80 per cent were designated as '?'. There are five exceptions to the rule: two systems (03571+3023 and 14162+3234) have $C \approx 0.77$ but were ruled out. Six systems (15015+2400, 16389+3643B, 21302+2312A, 21409+1824, 22085+1425A and 22085+1425B) have $C > 0.8$ but the predicted event was close to the beginning or end of an observing window, a possible event was observed significantly before the predicted time or different observations had conflicting results: these are designated as '?'.

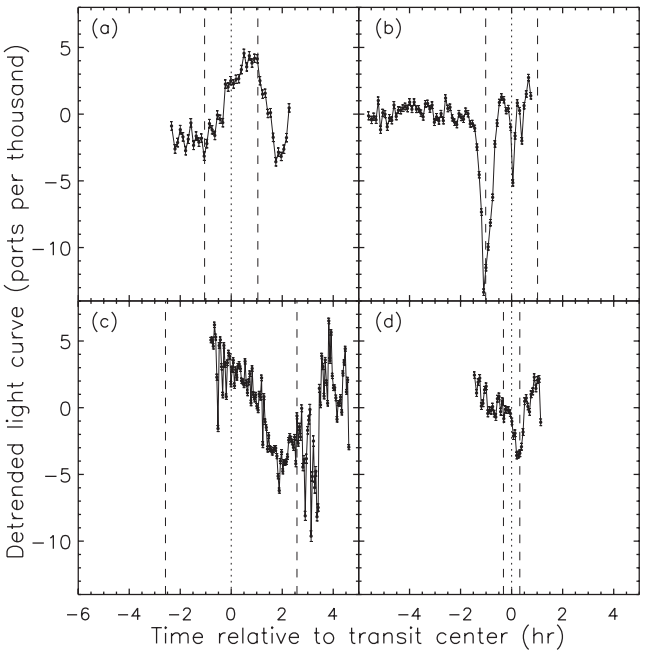

Figure 5. Detrended light curves from follow-up observations of four stars containing a transit-like event. The error bars show the 1σ errors from Poisson noise only. The vertical dotted lines mark the predicted transit time and the vertical dashed lines mark ± one standard deviation. The stars and UT epochs are (a) 03571+3023 on 2012 September 16, (b) 16442+3455 on 2013 May 3, (c) 17378+2257 on 2013 April 24 and (d) 18075+4402 on 2013 April 27.

3.2 Transit candidates

Follow-up observations of four signals produced light curves that contain a transit-like signal: 03571+3023, 16442+3455, 17378+2257 and 18075+4402 (Fig. 5). We have continued to observe predicted events for these stars to verify or rule out possible transits. 03571+3023 is variable: light curves are not consistent between predicted events and one is flat, causing us to rule out this system. 16442+3455 (Ross 813) was mis-assigned to a white dwarf in the catalogue of McCook & Sion (1987) but its colours and luminosity are clearly those of a late K or early M dwarf. 17378+2257 (GJ 686.1AB, HIP 86282) consists of a pair of dwarfs that have $V - J \approx 3$ and are designated as M0 stars in the Gliese catalogue

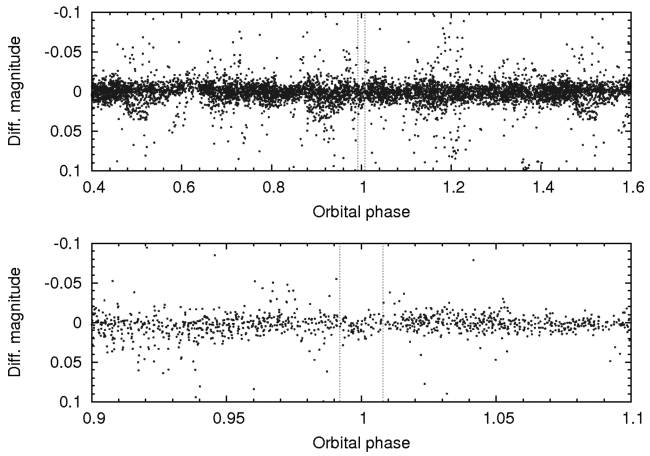

Figure 6. WASP light curve of GJ 436, which hosts a hot Neptune on a 2.64 d orbit. The data have been phased according to the established ephemeris of the planet and the transit is marked by the vertical lines. The bottom panel plots on an expanded scale.

but listed as K5 in Reid, Hawley & Gizis (1995). The molecular indices reported by Reid et al. (1995) and a spectrum obtained by us with the Mark III spectrograph on the MDM 1.3 m McGraw-Hill telescope suggest a spectral type between K7 and M0. These stars are an X-ray source (Hünsch et al. 1999) and the S-index values of the Ca II HK lines in their spectra (Duncan et al. 1991) suggest the stars are comparatively active (Isaacson & Fischer 2010), but Hα is not observed in emission (Young et al. 1989).

3.3 GJ 436b

GJ 436 aka SUPERBLINK star PM I11421+2642 or WASP source J114210.54+264230.4 is in our sample. The transit signal from its $4.3R_{\oplus}$ planet (Gillon et al. 2007) was not detected by the WASP pipeline and in fact no candidate signals were identified from this star. One possible explanation for the system's omission is that only one season of WASP data was obtained and the star fell 2.°4 from the centre of the field of view and thus was vignetted. Another contributing factor is the planet's high transit impact parameter ($b = 0.85$), which makes the transit unusually short. The data are also exceptionally noisy: the 5σ-filtered rms is 1.2 per cent, consistent with the nominal photometric error of 1.4 per cent and about 2.5 times the typical value for a $V = 10.7$ star (see equation 1). The transit is only marginally apparent even after the data are correctly phase-folded (Fig. 6). Although transits of GJ 436b were not detected by WASP, the inclusion of this system in our sample raises the question of whether we should expect additional hot Neptunes or whether this is the only such transiting system. For these reasons we carry out our statistical analysis for zero and one detections (Section 4).

4 ANALYSIS

4.1 Estimation of WASP detection limits

To place statistical constraints on the occurrence of hot Neptunes we calculated (i) the ability of HUNTER to detect planets in WASP light curves as a function of planet radius and orbital period, and (ii) the completeness with which our follow-up observations can rule out candidate transit signals (Section 4.2). Our criteria for WASP/HUNTER detection are the same as that applied to the data: SRN > 6, or SRN > 3 and signal-to-*white* noise $>\sqrt{50}$.

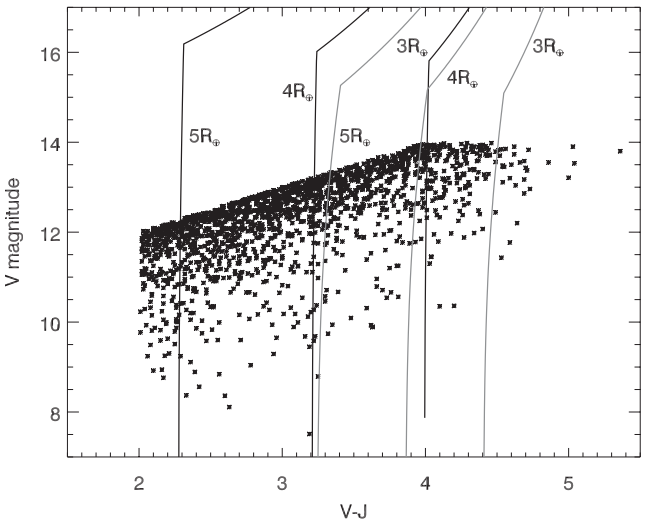

Figure 7. Expected WASP detection limits for Neptune-sized planets around late K and M dwarf stars, plotted versus $V - J$ colour (a proxy for $T_{\rm eff}$ and spectral type) and V magnitude. Transiting planets of specified radius (3, 4 or 5 $R_{\oplus}$), and orbital period (1.2 d, black curves, or 10 d, grey curves) should be detectable around stars to the right and below each curve. The stars of the SEAWOLF survey are plotted. Circular orbits, an impact parameter of zero and the median number of observations in our survey sample (8160) are assumed for these calculations.

The transit signal $\delta = (R_{\rm p}/R_*)^2$, where $R_{\rm p}$ is the radius of the planet. The red-noise error in the mean of N observations in the transit interval is $\sigma_1 N^{-\gamma}$, where σ_1 is the error in a single WASP measurement of a given star and the index $\gamma \approx 0.5$–$0.05 \times (15 - V)$ (based on fig. 2 in Collier Cameron et al. 2006). The white-noise error is taken to be $\sigma_1 N^{-0.5}$. We constructed an empirical formula for σ_1 based on fig. 2 in Collier Cameron et al. (2006):

$$\sigma_1 = 2.5 \times 10^{-3} \sqrt{1 + 8 \times 10^{(V-13)/5}}. \quad (1)$$

Assuming near-circular orbits for these close-in planets, the mean number of observations falling within a transit is taken to be $N = n\tau/P$, where τ is the duration of the transit and n is the total number of observations. The transit duration in days is

$$\tau \approx 0.075 R_* P^{1/3} \sqrt{1 - b^2}, \quad (2)$$

where R_* is in solar units, P is in days and b is the impact parameter (taken to be zero here).

Fig. 7 plots the limiting V magnitude for detecting a transiting Neptune-sized (3, 4 or $5R_{\oplus}$) planet around a dwarf star in the WASP survey as a function of stellar $V - J$ colour, our proxy for $T_{\rm eff}$ and stellar radius on the main sequence, for P =1.2 d or 10 d (see Section 4.3 for a justification for this range). The $V - J$ and V of stars in our survey catalogue are overplotted. Fig. 7 shows that WASP should be able to detect planets somewhat larger than Neptune (5–$6R_{\oplus}$) around nearly all of the stars in our sample, and planets slightly smaller than Neptune ($\sim 3R_{\oplus}$) around the coolest (M dwarf) stars, but only if they orbit quite close to their host. For orbital periods of 10 d only the largest Neptunes will be detectable around the M dwarfs.

The break in the slope of the detection contours at $V \approx 15$–16 in Fig. 7 is a result of a transition from a photon- or counting noise-limited regime to a red noise-dominated regime. Among stars with a fixed radius ($V - J$ colour) and $V >$ 15–16, detection improves with brightness. However, for $V <$ 15–16 correlated noise becomes important (decreasing γ with brighter V). For a fixed N, this means

that detection requires a lower σ_1 and hence an even brighter V. This positive feedback means that for a given stellar radius, planets below a certain size cannot be detected, regardless of apparent magnitude. This is a widely appreciated limitation of ground-based surveys (Pont et al. 2006; Smith et al. 2007).

4.2 Estimation of follow-up completeness

To evaluate the significance of non-detection or detection of transits in our follow-up observations, we calculated the completeness, i.e. the probability that a transit with the characteristics of the WASP candidate would be detected, and the false-alarm probability (FAP), i.e. the probability that such an event would be erroneously identified in our data in the absence of an actual transit.

Each follow-up observation of a candidate transit event yielded a normalized, de-trended light curve, plus errors based purely on counting statistics (Poisson or photon noise). FAP and completeness were calculated by constructing two sets of Monte Carlo realizations of the data, the first set with no transit signal added, and the second containing a transit signal equal in depth to the WASP candidate. The first set was used to set the detection threshold, i.e. the transit depth corresponding to a FAP of 0.01. This means there is a 1 per cent probability that a signal exceeding this threshold would be erroneously discovered in a light curve with these noise properties but containing no transit. We used the second set of Monte Carlo realizations plus the detection threshold determined from the first set to estimate the completeness or the recovery rate in follow-up photometry of transit signals with the properties of the WASP candidate.

To account for the effect of correlated or 'red' noise on transit detection, we computed the discrete autocorrelation function A_k of the actual data and used this to construct artificial light curves s_i of pure noise:

$$s_i = s \sum_j A_{i-j} w_j, \tag{3}$$

where w_i is a white-noise pattern and s is chosen so that s_i has the same total noise rms as the actual signal.

Our simple transit model used a linear limb-darkening law[5] and the 'small planet' approximation ($R_p \ll R_*$) such that the transit signal is

$$f(t) = \delta\,[1 - u\,(1 - \mu)]\,, \tag{4}$$

for $r < 1$, where $\mu = \sqrt{1 - r^2}$, u is the linear limb-darkening coefficient, the dimensionless radial coordinate is

$$r = \sqrt{(2(t - t_c)/\tau)^2 + b^2}, \tag{5}$$

and t_c is the transit centre time. Based on the median estimated $T_{\rm eff}$ of our sample (4570 K) and assuming solar metallicity, we adopted values of $u = 0.80$, 0.72 and 0.51 for Johnson V and R and Tiede J bandpasses, respectively (Claret 2000), and $u = 0.75$, 0.65 and 0.58 for Sloan riz bandpasses, respectively (Claret 2004). To calculate the transit duration τ we assumed a circular orbit and a stellar radius based on $V - J$ and Boyajian et al. (2012) (see Section 2.1). For each Monte Carlo realization, we drew a fixed value of impact parameter b from a uniform distribution limited to $\sqrt{3}/2$ (beyond which the transit duration is half the maximum value, resulting in exclusion from our sample).

[5] More complex limb-darkening laws are widely used but a linear law is completely adequate for creating and 'detecting' fake transits at low signal-to-noise ratio.

Table 3. Observations of candidate transits.

Star[a]	Observatory	t_c (MJD) (-2.45×10^6)	δ_d (10^{-3})	C
00177+2100	MDM	6191.85	16.6	0.032
00177+2100	MDM	6207.80	8.7	0.090
00492+2003	MDM	6195.75	2.3	0.644
01086+1714A	LCOGT/Faulkes	6212.81	1.4	0.940
01086+1714A	MDM	6284.71	6.4	0.206
01086+1714B	LCOGT/Faulkes	5898.79	7.5	0.047
01086+1714B	LCOGT/Faulkes	6157.03	3.9	0.360
01086+1714B	LCOGT/Faulkes	6564.98	7.2	0.061
01086+1714B	MDM	6583.69	5.1	0.165
01550+4035	LCOGT/BOS	6524.85	1.6	0.199

[a] ABC refer to multiple signals for the same star.
Table 3 is published in its entirety as a machine-readable table in the online version of this article and the Centre de Donées Strasbourg (CDS). A portion is shown here for guidance regarding its form and content.

To generate a distribution of false-positive transits, we fit the transit model to each transit-free light curve using the non-linear least-squares routine MPFIT (Markwardt 2009), with t_c and δ as free parameters. For the fit, an initial value of δ was chosen from a uniform distribution between 0 and 0.002. An initial value of t_c was chosen from a normal distribution with a standard deviation equal to the transit prediction error, and limited to the observation window. Cases where the fitted depth was negative or the transit was more than two standard deviations from the predicted transit centre were not counted, as these would have been excluded from the actual survey. We determined the 99 percentile value of the transit depth, corresponding to a FAP of 0.01. This is our adopted detection limit δ_d. This value was converted to an equivalent planet radius using the stellar radius, and we also computed a corresponding SNR detection threshold based on the white-noise rms of the light curve.

To calculate the completeness, we added artificial transits to the noise-only light curves and attempted to recover them. Each transit was modelled as described above, using the WASP candidate transit depth, a uniform distribution for b between 0 and $\sqrt{3}/2$ and a normal distribution for t_c. We then repeated the fitting process described above. To initially 'detect' the transit, we smoothed each light curve with a boxcar filter having a width equal to the expected transit duration. The minimum of this light curve became the initial guess at t_c in a fit with MPFIT. We calculated the fraction of recovered transit depths that exceeded δ_d, rejecting fits with t_c deviating from the actual value by more than two standard deviations. This fraction is our estimated completeness C. Table 3 reports values of δ_d and C for each follow-up observation.

4.3 Planet occurrence

Our observations constrain the intrinsic occurrence f (planets per star, or in the limit of few planets, fraction of stars with planets) of close-in Neptune- to Saturn-sized planets around late K and early M dwarf stars in the solar neighbourhood. A standard procedure to estimate f is to maximize a likelihood function that is the product of the probabilities of detections and non-detections. Our multi-stage observational campaign required us to consider how we defined detections and non-detections. Specifically, our sample includes the following.

(i) Stars with no transit-like signal found in WASP data: these were counted as non-detections.

(ii) Stars with a transit-like signal identified in WASP data but which were not screened with follow-up observations: these were considered as unconfirmed detections.

(iii) Stars with transit-like signals in WASP data which our follow-up observations have ruled out as transit candidates with some completeness C: these are considered possible non-detections or unconfirmed detections.

(iv) Stars with WASP signals that our follow-up photometry indicate are viable transit candidates: given sufficient follow-up, these could become confirmed detections.

Following Gaidos et al. (2013) we generalized the likelihood formalism as an empirical Bayes/marginalized likelihood analysis in which the occurrence rate f is a 'hyperparameter' of the prior probability that a star hosts a detectable transiting planet.[6] This prior is $\langle d_i \rangle f$, where $\langle d_i \rangle$ is the probability that a planet transits and is detected with the criteria in Section 2.4, marginalized over the distributions of planet radius and orbital period. The log-likelihood is

$$\ln \mathcal{L} = \sum_i^{\mathrm{ND}} \ln\left(1 - f\langle d_i \rangle\right) + \sum_k^{CD} \ln\left[(1 - F_k) f d_k\right] + \sum_j^{\mathrm{UD}} \ln\left[(1 - C_j) f d_j + C_j \left(1 - f\langle d_j \rangle\right)\right], \quad (6)$$

where the summations are over non-detections (ND), confirmed detections (CD), and unconfirmed detections (UD), C_j is the completeness of the follow-up observations that do not find a transit, and F_k is the FAP for detections confirmed by our follow-up observations. In the case of multiple observations of the same system we adopt the largest value of C_j.

If $fd \ll 1$ and $F \ll 1$, then

$$\ln \mathcal{L} \approx N_{\mathrm{CD}} \ln f - f \sum_i^{\mathrm{ND}} \langle d_i \rangle + \sum_j^{\mathrm{UD}} \ln\left[(1 - C_j) f d_j + C_j \left(1 - f\langle d_j \rangle\right)\right], \quad (7)$$

where N_{CD} is the number of confirmed detections. If C_i is not small for all stars (not the case here) then this can be further approximated as

$$\ln \mathcal{L} \approx N_{\mathrm{CD}} \ln f + \sum_j^{\mathrm{UD}} \ln C_j - f\left[\sum_i^{\mathrm{ND}} \langle d_i \rangle + \sum_j^{\mathrm{UD}} \left(C_j \langle d_j \rangle - \frac{1 - C_j}{C_j} d_j\right)\right]. \quad (8)$$

Only the first two terms depend on f, and from these one readily derives the most likely value

$$f_* = N_{\mathrm{CD}} \left[\sum_i^{\mathrm{ND}} \langle d_i \rangle + \sum_j^{\mathrm{UD}} \left(C_j \langle d_j \rangle - \frac{1 - C_j}{C_j} d_j\right)\right]^{-1}. \quad (9)$$

If no transits are confirmed the most likely value of f is zero.

The detection probability d_i for a given candidate transit signal is the product of a geometric factor d_i^{geo} and a detection probability d_i^{det}. Assuming circular orbits,

$$d_i^{\mathrm{geo}} \approx 0.238 R_* M_*^{-1/3} P^{-2/3}, \quad (10)$$

where the stellar parameters are in solar units and P is the signal period in days. d_i^{det} is estimated by computing both the SRN and $\Delta\chi^2$ for the given δ, P, and a uniformly distributed range of impact parameters, and determining the fraction of these which satisfy our selection criteria (Section 2.4). The SRN is given by $\delta N^{\gamma}/\sigma_1$, $\Delta\chi^2 = \delta^2 N/\sigma_1^2$, and σ_1, γ, and N are estimated as before. All cases with $b > \sqrt{3}/2$ were excluded because of the restriction on candidate transit duration (Section 2.4).

To calculate $\langle d \rangle$ we assumed a power-law distribution over radius $R_{\mathrm{min}} < R_{\mathrm{p}} < R_{\mathrm{max}}$ with index α, and a flat log distribution with orbital period $P_{\mathrm{min}} < P < P_{\mathrm{max}}$ (Cumming et al. 2008; Howard et al. 2012), i.e.

$$dN = \frac{f \left(R_{\mathrm{p}}/R_{\mathrm{min}}\right)^{-\alpha} \mathrm{d}\ln R_{\mathrm{p}}\, \mathrm{d}\ln P}{\alpha \ln\left(P_{\mathrm{max}}/P_{\mathrm{min}}\right)} \quad (11)$$

where $R_{\mathrm{p}} > R_{\mathrm{min}}$ and $P_{\mathrm{min}} < P < P_{\mathrm{max}}$. We calculated the minimum detectable planet radius R_{det} and the fraction of planets that would be detected, i.e.

$$d^{\mathrm{det}} = \begin{cases} \frac{R_{\mathrm{det}}^{-\alpha} - R_{\mathrm{max}}^{-\alpha}}{R_{\mathrm{min}}^{-\alpha} - R_{\mathrm{max}}^{-\alpha}}, & \text{if } R_{\mathrm{min}} < R_{\mathrm{det}} < R_{\mathrm{max}} \\ 1, & \text{if } R_{\mathrm{det}} < R_{\mathrm{min}} \\ 0, & \text{if } R_{\mathrm{det}} > R_{\mathrm{max}}. \end{cases} \quad (11b)$$

We marginalized over P and b assuming logarithmic and uniform distributions, respectively, and excluding values of either parameter that were also excluded during our selection of candidate transit signals, i.e. $b > \sqrt{3}/2$ and $P < 1.1$ d or periods within 5 per cent of artefacts (Section 2.4).

We adopted $R_{\mathrm{min}} = 3R_{\oplus}$ and $R_{\mathrm{max}} = 8R_{\oplus}$, i.e. slightly smaller than Neptune and Saturn, respectively. We adopted $P_{\mathrm{min}} = 1.2$ d and $P_{\mathrm{max}} = 10$ d following Howard et al. (2012) and Fressin et al. (2013). We determined that among 1728 stars with no detected signals within the range of $1.2 < P < 10$ and $3 < R_{\mathrm{p}} < 8$, assuming $\alpha = 1.9$, then $\sum_i^{\mathrm{ND}} \langle d_i \rangle = 16.8$. This is the expected number of detections around these stars if each had one such planet. It is not sensitive to the precise value of α.

We calculated the likelihood versus occurrence rate using equations (7) and (11), and (11b). Excluding GJ 436b, and given that we have as yet no confirmed detections of new planets in our sample, we can only place an upper limit on the occurrence of hot Neptunes. In this case, we place a 95 per cent confidence upper limit of 10.2 per cent on f based on a log likelihood within 1.92 of the maximum value (Fig. 8). We also estimate the most likely occurrence in the case of a single confirmed detection: $f = 5.3 \pm 4.4$ per cent (Fig. 8). The error is based on the assumption of asymptopic normality; a parabola was iteratively fitted to the log-likelihood curve and $\sigma_f = 1/\sqrt{2c}$, where c is the curvature of the parabola.

If there is more than one confirmed planet in our sample, the maximum likelihood estimate of f will be likewise higher. If we relaxed the assumption that all unconfirmed detections are ruled out, then f could be significantly higher, and close to unity, because the number of WASP candidates that we have yet to screen is comparable to the expected number (~17) if every star had a hot Neptune. However, if our follow-up results are representative of the results as a whole, then it is more likely that all or nearly all of these unscreened systems will be ruled out as well.

[6] Strictly speaking, the fraction of stars with such planets, which is equal to the number of such planets per star if the possibility of multiple planets in this restricted range of radii and orbital periods is neglected.

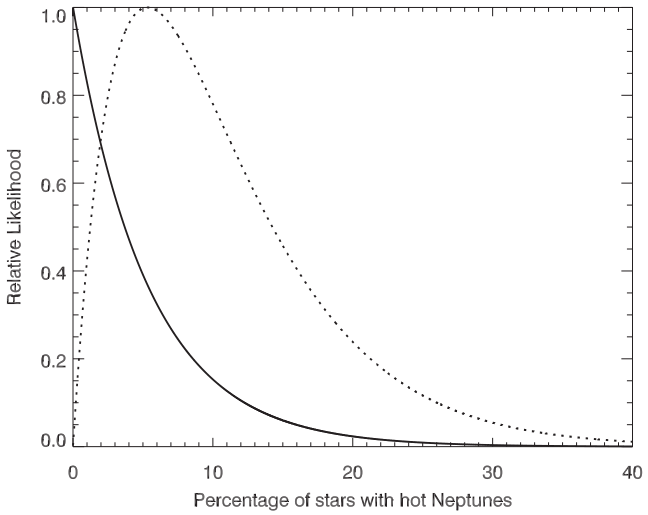

Figure 8. Likelihood versus occurrence of planets with 1.2 d $< P <$ 10 d and $3R_\oplus < R_\mathrm{p} < 8R_\oplus$ around SEAWOLF stars. The dashed line is for the case of one confirmed detection and the solid line is for the case of no confirmed detections.

4.4 Comparison with *Kepler*

We estimated the occurrence of 3–$8R_\oplus$ and $P <$ 10 d planets around *Kepler* target stars using the 2013 January release of confirmed and candidate transiting planets (KOIs) from analysis of observation quarters Q1–Q8. The methods are described in Gaidos et al. (2013) and here we recapitulate only the most important details. To emulate the range of spectral types of the SEAWOLF survey, stars with $2 < V - J < 4.7$, with V magnitudes based on the relation $V = r + 0.44(g - r) - 0.02$ (Fukugita et al. 1996), were selected from the complete *Kepler* target catalogue. We also required $K_\mathrm{p} < 16$ and that each star was observed for at least seven of the first eight observing quarters. We estimated parameters for these 14 578 stars and 190 (candidate) planets by fitting Dartmouth stellar evolution models (Dotter et al. 2008) using the Bayesian procedure described in Gaidos (2013). We then limited the analysis to 6422 stars with estimated $\log g > 4$ and g-D51 < 0.23. D51 is an AB magnitude based on a passband centred on 510 nm and the g-D51 colour is an indicator of gravity among K dwarfs; the colour-cut eliminates K giants (Brown et al. 2011). The median estimated T_eff of these stars is 4330 K. These stars host 136 candidate planets. Two of these have $3R_\oplus < R_\mathrm{p} < 8R_\oplus$ and $P <$ 10 d: KOIs 875.01 and 956.01 with R_p of 3.7 and $3.2R_\oplus$ on 4.22 and 8.36 d orbits, respectively.

We calculated the binomial log likelihood as a function of planet-hosting fraction f assuming a log distribution with orbital period and a power-law radius distribution in the limit that the transit probability is low (Mann et al. 2012; Gaidos et al. 2013):

$$f = \frac{N_\mathrm{p}\left(R_1^{-\alpha} - R_2^{-\alpha}\right)\ln\left(P_2/P_1\right)}{\alpha\sum_i^{N_D}\langle d_i\rangle} \tag{12}$$

where N_p is the number of detected planets with $R_1 < R_\mathrm{p} < R_2$ and $P_1 < P < P_2$, the summation in the denominator is over non-detections,

$$\langle d_i\rangle = \int_{R_1}^{R_2}\int_{P_1}^{P_2} R_\mathrm{p}^{-\alpha} d_i(R_\mathrm{p}, P)\,\mathrm{d}\ln P\,\mathrm{d}\ln R_\mathrm{p}, \tag{13}$$

and $d_i(R_\mathrm{p}, P)$ is the probability of detecting a planet around the ith star (Mann et al. 2012). For consistency with SEAWOLF we use $P_1 = 1.2$ d, $P_2 = 10$ d and $\alpha = 1.9$.

The transit of a late K or early M dwarf by a Neptune-sized planet produces a signal of magnitude 4×10^{-3}, far larger than the noise: the median 3 h Combined Differential Photometry Precision (CDPP3) for the stars in our sample is 1.8×10^{-4} and the 99 percentile value is 6.6×10^{-4}. We estimated the cumulative SNR from a $3R_\oplus$ planet on a 10 d orbit monitored for 2 yr (eight quarters): this is the least detectable case. The stellar noise over the transit interval was taken to be the CDPP3 scaled by $\sqrt{3/\tau}$ where τ is the transit duration in hours. Fressin et al. (2013) found that the recovery rate of the *Kepler* detection pipeline is nearly 100 per cent for SNR>16. Of the 9741 stars with CDPP3 values, for only 33 (0.3 per cent) would the estimated SNR be <16.

Thus the detection probability is essentially the geometric factor R_*/a, where a is the orbital semimajor axis, and independent of R_p. The R_p terms in equation (12) cancel and

$$f = N_\mathrm{p}\ln(P_2/P_1)/\sum_i^{\mathrm{ND}} F_i, \tag{14}$$

where F_i is defined by eqn. 5 in Gaidos et al. (2013).

The detection probability becomes

$$d_i(P) = \left(\frac{4\pi^2 R_*^3}{GM_*}\right)^{1/3}\frac{1 + e\cos\omega}{1 - e^2}P^{-2/3}, \tag{15}$$

where e is the orbital eccentricity and ω the longitude of periastron. Marginalizing over e and ω with an eccentricity distribution $n(e)$, and ignoring terms that do not depend on f,

$$\ln\mathcal{L} \approx N_D\ln f - 0.356 f\left[\int_0^1\frac{n(e)\mathrm{d}e}{1 - e^2}\right]\left(\frac{P_2}{1\mathrm{d}}\right)^{-2/3} \times\frac{(P_2/P_1)^{2/3} - 1}{\ln(P_2/P_1)}\sum_j^{N_D}\left(\frac{\rho_j}{\rho_\odot}\right)^{-1/3} + \cdots, \tag{16}$$

where N_D is the number of detected planets and ρ is the mean stellar density. Adopting the function for $n(e)$ in Shen & Turner (2008), we found that the integral is only weakly dependent on the parameter a in their distribution, and is ≈1.20 for $a = 4$. Using a Rayleigh distribution like that in Gaidos et al. (2013) gives a similar value of 1.08 for the integral. Because each star can be explained by more than one stellar model with probability p, we used a weighted mean of $\rho^{-1/3}$ to calculate the likelihood:

$$\langle\rho^{-1/3}\rangle = \sum_i p_i\rho_i^{-1/3}/\sum_i p_i, \tag{17}$$

where the summation is restricted to main sequence models, i.e. $\log g > 4$.

Under these assumptions, we found that the occurrence of hot ($P <$ 10 d) Neptunes is 0.33 ± 0.21 per cent (Fig. 9).

5 DISCUSSION AND CONCLUSIONS

We place a limit of 10 per cent on the occurrence of hot Neptunes ($P <$ 10 d) around the late K and early M dwarfs in our SEAWOLF sample (95 per cent confidence). In the event that a single planet candidate is confirmed, our maximum likelihood estimate of occurrence is 5.3 ± 4.4 per cent. From a Doppler survey of late F to early K stars, Howard et al. (2010) estimated an occurrence rate of about 8.1 ± 4 per cent for planets with projected masses $M\sin i$ of 10-100$M_\oplus$, a mass range correspondingly approximately to our radius

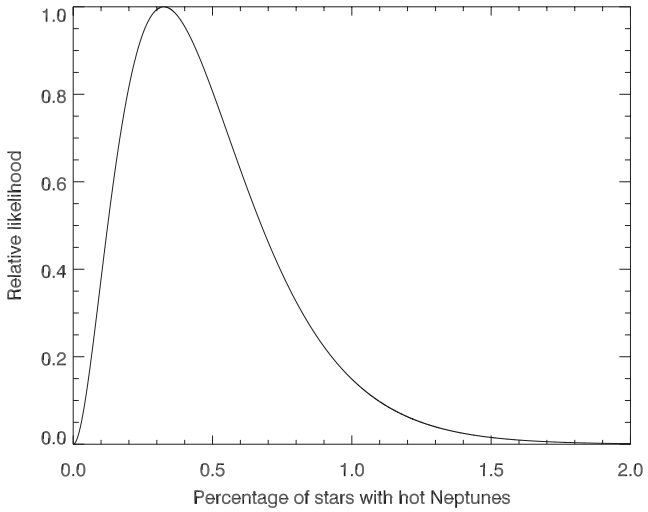

Figure 9. Likelihood versus occurrence of planets with $P < 10$ d and $3R_\oplus < R_p < 8R_\oplus$ around 6422 dwarf stars with $2 < V - J < 4.7$ observed by *Kepler* during at least seven quarters of Q1–8.

range, and $P < 50$ d. Assuming a logarithmic distribution with orbital period, and correcting by the factor $\ln(10/1.2)/\ln(50/1.2)$, the equivalent occurrence within 10 d is 4.6 per cent, a value similar to our estimate for the case of a single detection. Based on the HARPS Doppler survey, Mayor et al. (2011) estimated 11.1 ± 2.4 per cent of solar-type stars have planets with 10–30$M_\oplus$ within $P < 50$ d, but only 1.17 ± 0.52 per cent with masses of 30–100$M_\oplus$. Likewise, Bonfils et al. (2013) estimate that 3^{+4}_{-1} per cent of M dwarfs have 10–100$M_\oplus$ planets. That these Doppler-based values are consistent with the 5.3 per cent occurrence we derive assuming a single SEAWOLF detection suggests that GJ 436b may be the only transiting hot Neptune in our sample.

We estimated the occurrence of hot Neptunes around the late K and early M dwarf stars observed by *Kepler* to be 0.32 ± 0.21 per cent, more than an order of magnitude lower than in the SEAWOLF catalogue. Howard et al. (2012) report that the occurrence of 4–8$R_\oplus$ planets with $P < 10$ d around *Kepler* GK dwarfs is 0.23 ± 0.03 per cent, consistent with our estimate to within the errors (but for a different range of spectral types). One major caveat with interpreting the *Kepler* statistics is that late K spectral types also include red giant branch as well as dwarf stars; these luminosity classes can be difficult to distinguish by photometric colours alone and in the absence of spectroscopic screening, Malmquist bias will favour the inclusion of the large, more luminous stars (Gaidos & Mann 2013). Planets will be more difficult to detect and will appear smaller around RGB stars, e.g. some 'Earths' may actually be Neptunes. For this reason, it is possible that the statistical analysis of the *Kepler* results grossly underestimates the occurrence of hot Neptunes. However, our use of the *g*-D51 gravity-sensitive colour in constructing our *Kepler* sample should limit this effect. More spectroscopy of *Kepler* targets in this range of $V - J$ colours is needed to quantify contamination by RGB stars.

Our determination of an order-of-magnitude lower relative occurrence of hot Neptunes around *Kepler* stars echoes the findings of Wright et al. (2012), who found a deficit of hot Jupiters around these stars. One intriguing possibility is that *Kepler* stars are older, more evolved, and have more massive convective envelopes than those in the solar neighbourhood, and that close-in giant planets have suffered tidally driven decay of their orbits and been destroyed (Gaidos & Mann 2013). However, in the regime where the orbital period P is much shorter than the eddy turnover time-scale T, the rate of orbital decay is (Kunitomo et al. 2011)

$$\frac{\dot{a}}{a} = \frac{3}{4}\frac{M_p}{M_*}\frac{L_* P^2}{M_*(R_* - R_{env})^2}\left(\frac{R_*}{a}\right)^8, \quad (18)$$

where M_p is the mass of the planet, L_* the stellar luminosity, R_{env} the inner radius of the convective envelope, and a the orbital semimajor axis of the planet. For a Neptune-mass planet on a 10 d orbit around an M0 dwarf star the theoretical orbital decay time is $>10^{18}$ yr. The lower luminosity and smaller radius of K/M dwarfs relative to solar-type stars, and the lower mass of Neptunes relative to Jupiters, means that this process is too slow to explain the deficit of Neptunes close to *Kepler* stars relative to the solar neighbourhood. *Kepler* stars may also be more metal-poor than the solar neighbourhood, but there is, as yet, no evidence that the occurrence of Neptunes depends on the metallicity of the host star (Mann et al. 2013). Instead, the discrepancy must be explained by observational selection or differences in the efficiency in formation on or migration to close-in orbits.

The Next Generation Transit Survey (Wheatley et al. 2013) will monitor $\sim 1.6 \times 10^6$ K4-M4 dwarfs with $I < 17$ to search for hot Neptunes; about 2×10^5 of these stars will have $I < 15$ and are suitable for Doppler follow-up (P. Wheatley, personal communication). If our 0.32 per cent occurrence rate from *Kepler* is correct the number of transiting hot (super-)Neptunes in this survey will be ≤ 60, depending on actual detection efficiency. If the occurrence rate is close to 5 per cent, as suggested by Doppler surveys of nearby stars, the survey could find up to ~ 1000 such planets. These two values bracket an estimate by the NGTS team (Wheatley et al. 2013).

ACKNOWLEDGEMENTS

EG and KC acknowledge support from NASA grants NNX10AI90G (Astrobiology: Exobiology & Evolutionary Biology) and NNX11AC33G (Origins of Solar Systems). GM acknowledges support from the State Scholarships Foundation of Greece in the form of a scholarship and fruitful discussions with E. Paleologou. AZ acknowledges the Foundation for Research and Technology-Hellas. Space astrophysics in Crete is supported in part by EU REGPOT project number 206469. NN acknowledges support by an NAOJ Fellowship, NINS Program for Cross-Disciplinary Study, and Grant-in-Aid for Scientific Research (A) (No. 25247026) from the Ministry of Education, Culture, Sports, Science and Technology of Japan. EC and JB were supported by NASA/University of Hawaii Space Grant fellowships. The SuperWASP-N Camera was constructed and operated with funds made available from WASP Consortium universities and the UK Science and Technology Facilities Council (STFC). We extend our thanks to the Director and staff of the Isaac Newton Group of Telescopes for their support of SuperWASP-N operations. We thank the staff of the MDM and Skinakas observatories for their invaluable and courteous assistance during many observing runs. The Las Cumbres Observatory Network operates Faulkes Telescope North, ELP at McDonald Observatory, and the Byrne Observatory; the last is located on the Sedgwick Reserve, a part of the University of California Natural Reserve System.

REFERENCES

Bakos G. Á., Hartman J. D., Torres G., Kovács G., Noyes R. W., Latham D. W., Sasselov D. D., Béky B., 2011, in Bouchy F., Díaz R., Moutou C., eds, European Physical Journal Web of Conferences, Vol. 11,

Detection and Dynamics of Transiting Exoplanets. EDP Sciences, Les Ulis, p. 1002
Baraffe I., Chabrier G., Barman T. S., Selsis F., Allard F., Hauschildt P. H., 2005, A&A, 436, L47
Berta Z. K., Irwin J., Charbonneau D., Burke C. J., Falco E. E., 2012, AJ, 144, 145
Bonfils X. et al., 2012, A&A, 546, A27
Bonfils X. et al., 2013, A&A, 549, A109
Borucki W. J. et al., 2010, Science, 327, 977
Boué G., Figueira P., Correia A. C. M., Santos N. C., 2012, A&A, 537, L3
Boyajian T. S. et al., 2012, ApJ, 757, 112
Brown T. M., Latham D. W., Everett M. E., Esquerdo G. A., 2011, AJ, 142, 112
Brunini A., Cionco R. G., 2005, Icarus, 177, 264
Carone L. et al., 2012, A&A, 538, A112
Charbonneau D., Brown T., 2000, ApJ, 529, L45
Charbonneau D. et al., 2009, Nature, 462, 891
Christian D. J. et al., 2006, MNRAS, 372, 1117
Claret A., 2000, A&A, 363, 1081
Claret A., 2004, A&A, 428, 1001
Collier Cameron A. et al., 2006, MNRAS, 373, 799
Cox A. N., 2000, Allen's Astrophysical Quantities. Am. Inst. Phys., New York
Cumming A., Butler R. P., Marcy G. W., Vogt S. S., Wright J. T., Fischer D. A., 2008, PASP, 120, 531
Dotter A., Chaboyer B., Jevremović D., Kostov V., Baron E., Ferguson J. W., 2008, ApJS, 178, 89
Duncan D. K. et al., 1991, ApJS, 76, 383
Fressin F. et al., 2013, ApJ, 766, 81
Fukugita M., Ichikawa T., Gunn J. E., Doi M., Shimasaku K., Schneider D. P., 1996, AJ, 111, 1748
Gaidos E., 2013, ApJ, 770, 90
Gaidos E., Mann A. W., 2013, ApJ, 762, 41
Gaidos E., Fischer D. A., Mann A. W., Howard A. W., 2013, ApJ, 771, 18
Gillon M. et al., 2007, A&A, 472, L13
Hansen B. M. S., Murray N., 2012, ApJ, 751, 158
Hartman J. D. et al., 2011, ApJ, 728, 138
Henry G. W., Marcy G. W., Butler R. P., Vogt S. S., 2000, ApJ, 529, L41
Howard A. W. et al., 2010, Science, 330, 653
Howard A. W. et al., 2012, ApJS, 201, 15
Hünsch M., Schmitt J. H. M. M., Sterzik M. F., Voges W., 1999, A&AS, 135, 319
Isaacson H., Fischer D., 2010, ApJ, 725, 875
Kovács G., Zucker S., Mazeh T., 2002, A&A, 391, 369
Kovács G., Bakos G., Noyes R. W., 2005, MNRAS, 356, 557
Kunitomo M., Ikoma M., Sato B., Katsuta Y., Ida S., 2011, ApJ, 737, 66
Laughlin G., Bodenheimer P., Adams F. C., 2004, ApJ, 612, L73
Lépine S., Gaidos E., 2011, AJ, 142, 138
Lépine S., Shara M. M., 2005, AJ, 129, 1483
McCook G. P., Sion E. M., 1987, ApJS, 65, 603
McNeil D. S., Nelson R. P., 2010, MNRAS, 401, 1691
Mann A. W., Gaidos E., Aldering G., 2011, PASP, 123, 1273
Mann A. W., Gaidos E., Lépine S., Hilton E. J., 2012, ApJ, 753, 90
Mann A. W., Gaidos E., Kraus A., Hilton E. J., 2013, ApJ, 770, 43
Markwardt C. B., 2009, in Bohlender D. A., Durand D., Dowler P., eds, ASP Conf. Ser. Vol. 411, Astronomical Data Analysis Software and Systems XVIII. Astron. Soc. Pac., San Francisco, p. 251
Mayor M. et al., 2011, A&A, preprint (arXiv:1109.2497)
Mordasini C., Alibert Y., Benz W., 2009, A&A, 501, 1139
Pollacco D. L. et al., 2006, PASP, 118, 1407
Pont F., Zucker S., Queloz D., 2006, MNRAS, 373, 231
Reid I. N., Hawley S. L., Gizis J. E., 1995, AJ, 110, 1838
Rogers L. A., Bodenheimer P., Lissauer J. J., Seager S., 2011, ApJ, 738, 59
Shen Y., Turner E. L., 2008, ApJ, 685, 553
Skrutskie M. F. et al., 2006, AJ, 131, 1163
Smith A. M. S. et al., 2007, in Afonso C., Weldrake D., Henning T., eds, ASP Conf. Ser. Vol. 366, Transiting Extrapolar Planets Workshop. Astron. Soc. Pac., San Francisco, p. 152
Southworth J. et al., 2009, MNRAS, 396, 1023
Tamuz O., Mazeh T., Zucker S., 2005, MNRAS, 356, 1466
Wheatley P. J. et al., 2013, in Saglia R., ed., European Physical Journal Web of Conferences, Vol. 47, Hot Planets and Cool Stars. EDP Sciences, Les Ulis, p. 13002
Wright J., Marcy G., Howard A., Johnson J. A., Morton T., Fischer D. A., 2012, ApJ, 753, 160
Yanagisawa K. et al., 2006, SPIE Conf. Ser. Vol. 6269, Ground-based and Airborne Instrumentation for Astronomy. SPIE, Bellingham
Young A., Skumanich A., Stauffer J. R., Harlan E., Bopp B. W., 1989, ApJ, 344, 427

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Table 2. Candidate Transit Systems Identified in WASP Data.
Table 3. Observations of Candidate Transits (http://mnras.oxfordjournals.org/lookup/suppl/doi:10.1093/mnras/stt2078/-/DC1).

Please note: Oxford University Press are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

This paper has been typeset from a TEX/LATEX file prepared by the author.